The "Task B Problem" and Other Considerations in Developmental Functional Neuroimaging

Jessica A. Church,[1] Steven E. Petersen,[1,2,3,4] and Bradley L. Schlaggar[1,2,3,5]*

[1]*Department of Neurology, Washington University School of Medicine, St. Louis, Missouri*
[2]*Department of Radiology, Washington University School of Medicine, St. Louis, Missouri*
[3]*Department of Anatomy and Neurobiology, Washington University School of Medicine, St. Louis, Missouri*
[4]*Department of Psychology, Washington University in St. Louis, St. Louis, Missouri*
[5]*Department of Pediatrics, Washington University School of Medicine, St. Louis, Missouri*

Abstract: Functional neuroimaging provides a remarkable tool to allow us to study cognition across the lifespan and in special populations in a safe way. However, experimenters face a number of methodological issues, and these issues are particularly pertinent when imaging children. This brief article discusses assessing task performance, strategies for dealing with group performance differences, controlling for movement, statistical power, proper atlas registration, and data analysis strategies. In addition, there will be discussion of two other topics that have important implications for interpreting fMRI data: the question of whether functional neuroanatomical differences between adults and children are the consequence of putative developmental neurovascular differences, and the issue of interpreting negative blood oxygenation-level dependent (BOLD) signal change. *Hum Brain Mapp 31:852–862, 2010.* © 2010 Wiley-Liss, Inc.

Key words: fMRI; methods; negative BOLD; group differences

INTRODUCTION

Functional magnetic resonance imaging (fMRI) is a tool that allows scientists and clinicians unprecedented, noninvasive access to brain activity in children. Because fMRI has minimal risk, no known long-term effects, and does not involve radiation, researchers have an alternative to EEG and other surface-based tools to study both structural and functional brain development. In addition, fMRI's high spatial resolution and noninvasive nature makes it well suited for the study of cognition in children. While knowledge about the precise relationship of the fMRI signal to neuronal activation continues to expand [Logothetis et al., 2001; Raichle and Mintun, 2006], the utility of fMRI for clinical pediatric and developmental cognitive neuroscience is clear.

Despite the promise of functional neuroimaging to study cognitive development, researchers continue to face a number of methodological issues. The goal of the following discussion is to describe some of these issues, their implications, and ideas regarding how to address them. Greater consistency of methods and techniques across laboratories will allow for easier comparison of results and accelerate our understanding of functional brain development.

Contract grant sponsor: NIH; Contract grant numbers: K02NS053425, R01HD057076, R21NS061144; Contract grant sponsor: McDonnell Center for Systems Neuroscience.

*Correspondence to: Bradley L. Schlaggar, Department of Neurology, Box 8111, 4525 Scott Avenue, Washington University School of Medicine, St. Louis, Missouri 63110, USA.
E-mail: schlaggarb@neuro.wustl.edu

Received for publication 3 December 2009; Revised 20 January 2010; Accepted 24 January 2010

DOI: 10.1002/hbm.21036
Published online 27 April 2010 in Wiley InterScience (www.interscience.wiley.com).

The Performance Burden

Accounting for task performance

An important issue in developmental imaging, but also for any group comparison, is group differences in task

performance [Brown et al., 2005; Johnson et al., 2002; Palmer et al., 2004; Price and Friston, 1999; Schlaggar and McCandliss, 2007; Schlaggar et al., 2002]. A task that is simple for adults could be much harder for children, or a task designed for children could produce ceiling effects in adults. A discrepancy in performance on the task of interest (as well as any comparison task—see "the Task B Problem" later) creates a potential confound in the analysis. In this instance, any differences in activation observed between groups might be due to less successful performance (e.g., inattentiveness, misunderstanding of instructions, and guessing) in one group and not necessarily by a fundamental difference in the way the brains of members of the two groups process the task. This point is important, because if we want to discover group differences in brain processing responsible for producing a particular behavior, we should do our best to increase the chance that we are sampling the brain activity during that behavior. If group task performance is discrepant, there are a number of reasons beyond the fundamental group differences in brain processing related to task implementation, for why this discrepancy could be the case. In studies that do not address the potential confound of performance discrepancies, results must be interpreted with caution. To be clear, understanding the functional neuroanatomical basis of group performance differences is also valuable. The argument here is that by isolating the variables, one can attain the unique contribution of performance or group membership to any differences observed between groups.

The first step in addressing the performance confound is to collect, whenever possible, behavioral data while the subject is in the scanner (i.e., recordings of verbal outputs, eye movements, button presses, etc). Performance metrics should include both accuracy and reaction time. One can argue that response accuracy and response time are nonoptimal surrogates of performance (though chronometrics have certainly provided the means for interrogating cognitive architecture [Posner, 1978]). Two groups of subjects might be entirely matched on accuracy and response time, yet have two entirely different strategies, whether those strategies are overt or implicit, for task completion [Brown et al., 2005; Schlaggar et al., 2002]. In developmental (and aging) studies, neurobiological differences may bias toward different implicit processing strategies over age [Grossman et al., 2002; Reuter-Lorenz and Lustig, 2005]. Group imaging differences, then, could result from one group's unsuccessful implementation of the same or alternate strategy as the other group (i.e., "the performance confound"), or successful implementation of an alternate strategy (i.e., "behavioral phenocopy," see Interpretation of Group Differences below). When behavioral performance is not different between the groups, but imaging differences remain, the interpretation space is narrowed to consideration of successful, as measured by overt performance, implementation of different strategies. Without overt performance information, there is no means to address a frank performance confound.

Relative to a typical testing environment, the scanner can cause a degradation in performance in subjects of all ages, but children may be more susceptible than adults. Hence, estimates of an individual's performance on the task outside of the scanner cannot be relied upon exclusively to give a good estimate of their performance when the imaging data were acquired. Similarly, it is important to differentiate between correct and incorrect trials of a task, as it has been shown that error-related activity can differentially affect many regions of the brain [Dosenbach et al., 2006; Garavan et al., 2002], and that there may be group differences in error processing [Rubia et al., 2005; Velanova et al., 2008]. Examination of just the correct trials of a task does not address differences in response times between the two groups. There are well-established age differences in processing speed across a variety of tasks [Kail, 1991]. Thus, reaction time effects are important to distinguish from other types of group activation differences.

Research groups have dealt with performance differences between groups using many strategies. We will discuss four strategies briefly. One common technique is to create equivalent performance by calibrating the demands of the task until the adults and children are performing at a similar level of accuracy and/or reaction time [summarized briefly in Casey, 2002; Kotsoni et al., 2006]. Testing various levels of task difficulty in each of the groups allows for comparison of activity at equivalent performance. For example, one could parametrically manipulate an N-back working memory task to create roughly equal performance between children and adults (e.g., by having children do a 1-back version and adults do a 3-back version). However, this parametric manipulation assumes that the brain activations in the two groups are being manipulated the same way by the different versions of the task. This may be problematic in cases of memory span, for instance, when different list lengths are proposed to emphasize different processes.

Another strategy for addressing performance differences is that of post-hoc "performance matching." In this approach, the groups perform the same task, and any group differences are found. A subgroup analysis is then done by separating groups based on overt performance measures (i.e., reaction time and accuracy) into matched and unmatched sets. Using this approach, we have identified regions that produce group differences only when performance is discrepant between groups, and regions that remain different between groups even when performance is equated [Brown et al., 2005; see "Interpretation of Group Differences" later]. This approach requires some degree of overlapping performance on the task between the groups, which may not be possible for all tasks. This method has been criticized as selecting for the slowest and dullest adults and the quickest and brightest children in the subgroups. However, it is important to note that the group imaging differences continue to exist in the other subgroups [Brown et al., 2005; Schlaggar et al., 2002]. Also, task behavioral responses are often single "moment in

time" measures; subgroups, and the overall groups, often can be matched on IQ and other offline assessments suggesting the "bright/dull" dichotomy is not inherently built-in to this approach.

A third strategy is to regress performance variables as covariates of interest. Performance regression is often done in conjunction with an age regression, and requires at least some degree of non-collinearity between age and performance [see Fair et al., 2006 for discussion]. Performance regression has the benefit of not reducing power through subgroupings as performance matching analysis can do, but a strong degree of collinearity between age and performance may inflate the variance of the estimate related to each of the factors.

A fourth strategy for dealing with performance issues has been to equate performance between the groups on one task (e.g., picking out the capital letter in a word), while indirectly assessing the group differences on a simultaneously occurring implicit task (e.g., reading). However, this approach raises issues similar to the Task B problem discussed next, in that there is little reason to assume that because the groups are equated on the overt task that they are thus matched on the implicit task.

The "Task B" problem

Many neuroimaging studies employ direct comparisons between two (or more) task conditions, such that brain activity during a "control task" (Task B) is subtracted from brain activity during a "higher order" cognitive task (Task A). The thought is that an appropriate control task will subtract away functional activity common to the two tasks, leaving only activity related to the higher order aspects of Task A ("pure insertion"). The concept of pure insertion is problematic because of its assumptions of linearity and noninteraction between tasks [Friston et al., 1996]. The choice of a control task thus has important consequences on results and their interpretation, particularly in group comparison studies where Task B has the potential to be different between the groups. For example, developmental studies often compare (Task A_{child}–Task B_{child}) to (Task A_{adult}–Task B_{adult}), and interpret the results as a straightforward difference between children and adults for Task A. Critically, this interpretation rests on the assumption that Task B is the same in both children and adults, but this assumption may go untested or not be discussed. This assumption is very common in blocked designs. However, if the two groups activate significantly different brain regions (or activate similar regions to different degrees) while performing a control task, the interpretation of group differences in the higher order task will be confounded (see Fig. 1). This issue can be addressed in at least two complementary manners. First, one can directly compare the two groups on Task B (Task B_{adult}–Task B_{child}). A null result in this direct comparison would support the contention that Task B is behaving well as a comparison task. Some might worry that a "Task C problem" emerges here such that Task B needs to be contrasted with a lower level task, *ad infinitum*. The key is that the between-group Task B comparison obviates this need. The point of the between-group Task B comparison is to test the validity of the (Task A_{child}–Task B_{child}) versus (Task A_{adult}–Task B_{adult}) construct. Alternatively, or in addition, one can investigate a Task (A vs. B) by Group (child vs. adult) interaction and use post hoc analysis to determine the source of the interaction.

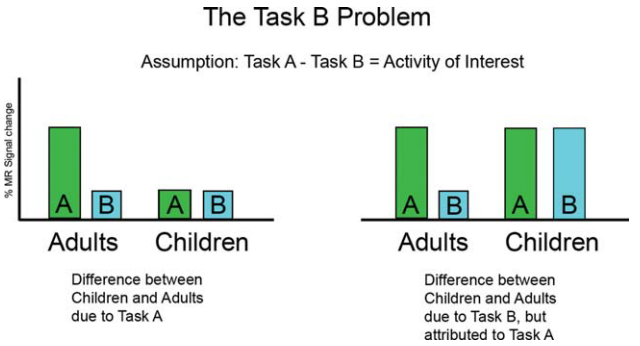

Figure 1.
The Task B Problem. Studies that compare the difference in activation between two tasks across two groups must account for the possibility of differences existing in either task (Task A or Task B), not just the task of interest (Task A).

Event-related fMRI designs allow estimation of the hemodynamic response to individual trial types, eliminating the need for subtraction and thus eliminating the Task B problem. When these designs do not assume a shape of response, timecourses of the hemodynamic responses can be evaluated and displayed. Examination of timecourses allows a straightforward quality control assessment to be implemented to ensure appropriate event coding and data preprocessing. Analysis of the timecourses of separate conditions also reveals the direction of hemodynamic response to an event type (positive or negative activity) that other types of analysis (particularly subtraction analyses) can obscure; the direction of the response can be critical to data interpretation (see "Negative BOLD Activity" below).

An approach that we often employ begins with the use of an event-related (or a mixed/blocked event-related) design to assess group differences. We model a parameter for every time point in the time course of each transient event type (e.g., correct vs. incorrect responses, different item types) thus estimating the magnitude at each of the points while making no assumptions about the precise shape of the response [Corbetta et al., 2000; Shulman et al., 1999]. From here, a significant main effect of time from an ANOVA analysis can be conducted at the voxel level or by using experimenter-defined regions. A significant effect indicates that the hemodynamic response deviates significantly from zero (i.e., is not flat), and can be interpreted as an activation or deactivation. A significant interaction of any other factor with time implies a significant variation

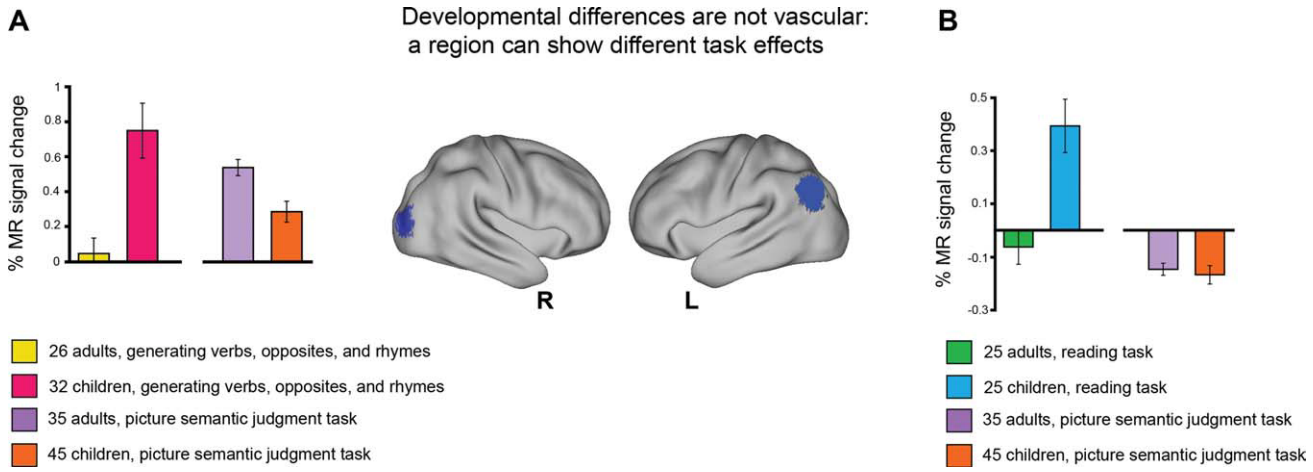

Figure 2.

Developmental differences are not vascular. This figure demonstrates the peak magnitude of BOLD activity observed for a right extrastriate region (**A**), and a left angular gyrus region (**B**), for two separate tasks. (A) In the right extrastriate region, a study of aloud verb, opposite, and rhyme generation to visually presented words produced greater activity in 32 children (ages 7–10 years) than in 26 adults (ages 18–32 years). However, this same region produced greater activity in 35 adults (ages 21–29 years) during a picture semantic judgment task compared with 45 children (ages 7–9 years). Showing a double dissociation of BOLD signal change between similarly-aged children and adults for different tasks is not readily explained by vascular effects. (B) In a high-frequency word reading task, 25 children (ages 7–10 years) have positive activity in a region of the left angular gyrus, while 25 adults (ages 18–32 years) do not significantly activate the region. However, when this region was applied to the groups who performed the picture judgment task (described for A), activity in this region was significantly negative for both groups. These demonstrably different directions of BOLD activity within similarly-aged children is not readily explained by vascular effects.

in the hemodynamic response across the levels of that factor. This approach allows for independent assessment of both Tasks A and B and their effect across the groups.

Group Differences Beyond Task Performance

The vasculature argument

One critique of developmental imaging is that the differences observed between children and adults are due to differences in neurovascular coupling that underlie the BOLD signal, and thus are artifacts. This concern, it seems, is motivated by the fact that the precise relationship between neural activity and BOLD signal is not well understood. BOLD is an indirect measure of neural activity, relying on local net increase in blood oxygenation to provide the contrast used in fMRI studies. If it turns out that the relationship between neuronal activity and BOLD signal generation has a developmental timecourse then, the argument goes, it cannot be ruled out that group differences are due to developmental changes in vascular physiology (as opposed to neural information processing).

This scenario seems highly unlikely based on at least four sets of observations. First, when children and adults are put in a common atlas space and perform a simple visual task, timecourses are similar in children and adults in regions located across three different vascular distributions [i.e., motor cortex, middle cerebral artery; supplemental motor area, anterior cerebral artery; and visual cortex, posterior cerebral artery; Kang et al., 2003]. Second, both increases and decreases in brain activity between age groups are often observed from performance of the same task [Bunge et al., 2002; Konrad et al., 2005; Velanova et al., 2009]. Increases and decreases in BOLD are seen even when there are no performance differences between groups, and, importantly, much of the brain BOLD activity is the same between groups [Brown et al., 2005; Church et al., 2008]. Third, different tasks can result in different relationships of activity between children and adults within the same region (see Fig. 2). In some cases, adult vs. child activation relationships go in opposite directions based on the task (e.g., BOLD activity in a region can be greater in children than adults during one task, but be greater in adults than children in a different task). Such manipulations of activity should not be possible if the differences in brain activity over age were due to neurovascular effects. Fourth, across the emerging developmental cognitive neuroscience literature, the location of group differences is task-dependent, not task-independent. A developmental difference in neurovascular coupling cannot account for this array of observations.

Head size and movement

Children have slightly smaller heads than adults [although 7-year-olds' brains are ~95% of adult size, and

have adult-level gyrification; Armstrong et al., 1995; Caviness et al., 1996; Giedd et al., 1999; Lenroot and Giedd, 2006]. Children also tend to move more in the scanner [Brown et al., 2005; Yuan et al., 2009]. Both of these issues can cause increased and potentially, systematic variability (i.e., colored noise) in the data. For the appropriate age groups, experimenters can minimize the impact of head size by putting adult and child data into a common space (see "Common Atlas" section below). Head size can also be measured, and its impact on results can be assessed using subgroupings based on head size. Similarly, grey matter density could be compared between groups using voxel-based morphometry [Ashburner and Friston, 2000].

Movement during scan acquisition can be dealt with using a number of techniques, including vacuum pillows, masks, and other head restraint systems [Burgund et al., 2002; Mathur et al., 2008; Raschle et al., 2009]. In addition to appropriate movement training and head movement restriction during scans, we measure movement in six different parameters (x, y, and z directions in millimeters, and degree of rotation around x, y, and z axes) on a frame-by-frame basis during all scan sessions. Typically, any participant who moves more than a given threshold is removed from analysis. We also check separately to ensure that no movement is strongly correlating with signal estimates. Other techniques include creation of movement-matched and nonmatched subgroups to show that the results are not due exclusively to movement, or regression of movement as a covariate of no interest in the initial single-subject general linear model.

Statistical power, variability, and success rate

Sufficient power to detect subtle differences between groups is critical to successful neuroimaging studies. Recent articles have highlighted the importance of large groups and independent replication to avoid erroneous results [Ihnen et al., 2009; Thirion et al., 2007]. Developmental studies should thus plan on recruiting relatively large numbers of subjects to achieve appropriate power [the calculation of which has recently been made easier by Mumford and Nichols, 2008], as well as striving for replication of results.

Experimenters conducting developmental studies are familiar with the lower successful scan rate in children compared to adults. The lower rate of success can be due to movement or performance difficulties as described above, or due to related incomplete scan sessions. We use a realignment protocol that allows children to leave the scanner for a break with minimal difficulties in acquisition, but some children, particularly those with developmental disorders, or ages younger than 7 years, are still not always successful at completing an entire scan session. This issue has been examined by Yerys et al. [2009], and their data suggests acquiring at least 10–15% more children than adults to have equal group acquisition. The number of extra subjects needed goes higher with any disorder or younger age [O'Shaughnessy et al., 2008; Yerys et al., 2009].

The likelihood of successful scan acquisition in children can be improved substantially by including a practice session with a "mock" scanner prior to the actual scan [O'Shaughnessy et al., 2008; Raschle et al., 2009]. Inclusion of movement training, perhaps using a motion sensor that interrupts a video the child is watching while in the mock scanner, might also help children understand the importance of holding still.

Optimizing and Interpreting Developmental Group Differences

Direct comparisons and common atlas space

Functional neuroimaging, as with any experimental approach, should always strive to be as objective and statistically valid as possible. Thus direct statistical comparisons of the data should be the gold standard of group comparisons. Side-by-side pictures of groups thresholded at the same level can be informative at a qualitative level, but no firm conclusions about group difference should be drawn without a probabilistic measure of the reliability of the similarities and differences in activation across groups.

Direct comparison generally requires the use of a common atlas or some other cross-subject registration. When developmental imaging was starting, it was unclear if child and adult brains were too dissimilar to share a common atlas. However, numerous studies have now had great success showing minimal warping differences when putting children (i.e., $\geq$7 years old) and adults into a common atlas [Burgund et al., 2002]. This approach allows the direct statistical comparison via standard analysis techniques. Alternatively, a direct comparison of anatomically-defined regions of interest could be done on an individual subject level without common atlas registration.

Voxel counting, where voxels (three-dimensional pixels) active over a given threshold form a region of interest and are counted for each group, however, fails to account for information about the variability and magnitude of the brain activity in each group. It thus does not take full advantage of the data available, and produces less than optimal statistics [e.g., see Cohen and DuBois, 1999].

Interpretation of group differences

When differences in performance and the other factors discussed above are accounted for, remaining group differences can imply different brain processing for a given task over age. As described by Schlaggar and McCandliss: "The presence of regions of the brain that show functional activation differences between adults and children, even when performance dynamics are matched, can be considered a behavioral phenocopy; the notion that identical performance is observed across groups yet is supported by different underlying neural mechanisms" [Schlaggar and McCandliss, 2007].

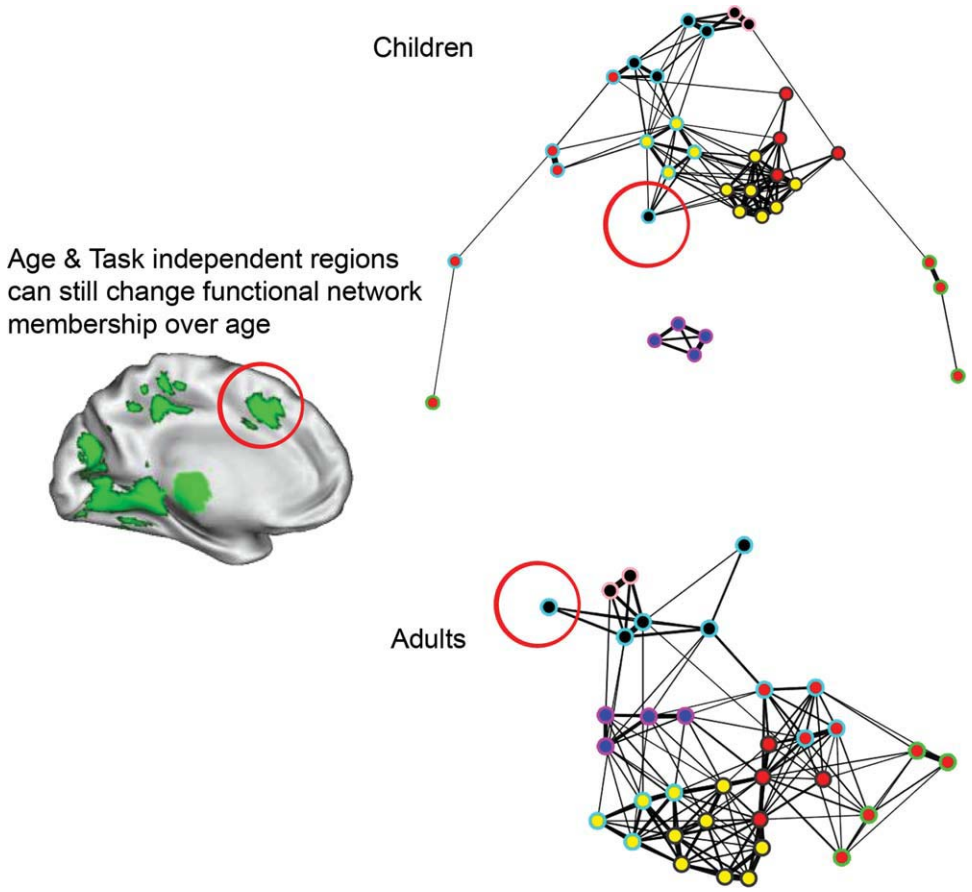

Figure 3.

Similar task activity is supported by different network relationships. A region in the dorsal anterior cingulate/medial superior frontal cortex was not significantly different between children (ages 7–10 years) and adults (ages 18–32 years) for controlled lexical processing tasks (green regions on brain) [Brown et al., 2005]. When this region (circled in red) was incorporated in a larger resting-state functional connectivity analysis, it was found to shift membership from a community of regions comprised of other frontal regions (blue outer rings of regions, top panel, 60 children ages 7–9 years), to being part of a putative task control network (black colored regions, bottom panel, 60 adults ages 22–31 years) [Fair et al., 2009]. Circles on the two graphs represent seed regions, with the center color reflecting mature network relationships, and the outer ring color reflecting anatomical location. Center colors: yellow is the fronto-parietal control network, black is the cingulo-opercular control network, red is the default network, and blue is the cerebellum. Outer ring colors: cyan is frontal, salmon is subcortical, gray is parietal, green is temporal, and magenta is cerebellum. Lines represent correlations between pairs of regions at a threshold greater than 0.10, where line thickness scales by the strength of correlation. See Fair et al. [2009], for the full description and a movie of the transitions of networks from child to adult.

Some might argue that activation differences in the absence of any behavioral difference are of little relevance. We take a complementary position; when performance discrepancies between children and adults are absent (or minimized), the investigator has an excellent opportunity to measure brain processing differences that are most likely related to maturational effects, independent of accuracy or speed. As Johnson et al. have argued, the precise structure–function organization of the brain is not static over development; functional interactions and relationships reorganize over age in both typical and atypical development. Behavioral phenocopy reveals the added benefit of functional neuroimaging toward elucidating the developmental changes of cognitive processes in typical and atypically developing populations [Johnson et al., 2002; Schlaggar and McCandliss, 2007].

An alternative technique: Resting-state functional connectivity

It is interesting to note at this juncture that a relatively new form of image analysis has arisen which allows

comparison of groups with minimal task performance demands [Biswal et al., 1995]. Resting-state functional connectivity MRI (rs-fcMRI) examines the low-frequency spontaneous BOLD signal fluctuations that occur when a subject is resting in the scanner for several minutes [Biswal et al., 1995; Cordes et al., 2000; Damoiseaux et al., 2006; Dosenbach et al., 2007; Fair et al., 2007b; Fox and Raichle, 2007; Fox et al., 2005; Lowe et al., 1998]. It is hypothesized that correlations in this signal between different regions of the brain reflect the regions' functional relatedness over time [Dosenbach et al., 2008; Fox and Raichle, 2007]. This analysis of the BOLD signal has been of particular interest to investigators interested in typical and atypical development because of the lack of task demands [Church et al., 2009; Fair et al., 2007a, 2008; Just et al., 2007; Kelly et al., 2007]. A typical analysis examines regions of interest and correlations of their BOLD signal timecourses with those of other regions. Using graph-theory metrics [Honey et al., 2007; Karrer et al., 2008], putative network relationships can be discovered and directly compared between groups [Church et al., 2009; Fair et al., 2007a, 2009]. As widespread interest in this type of analysis is relatively recent, many fundamental questions are open to exploration. Investigations are ongoing to explore the effects of physiological factors (e.g., heart rate and respiration) on the low-frequency BOLD signal, as well as how these factors might change between groups or over development [e.g., Chang et al., 2009; Jones et al., 2010; Thomason et al., 2005]. It is also possible that there are systematic group differences or developmental differences in the "performance" of quiet rest. Future investigations should clarify these questions. The analysis of rs-fcMRI data includes a rapidly developing set of techniques that illuminate developmental change in the functional network organization of brain regions. This additional information can generate new questions and help to contextualize the similarities and differences found in task data over development.

A particularly exciting finding revealed by rs-fcMRI studies is that even if no age or performance effects are demonstrable in an fMRI activation study, that this does not mean from a network level view that no developmental effect is taking place. For example, a region in dorsal anterior cingulate/medial superior frontal gyrus (dACC/msFC) was shown to have no significant differences due to age or performance in an fMRI study of reading-related tasks [Brown et al., 2005]. However, in a rs-fcMRI study of regions involved in two putative control networks, the same region is part of one control network in children, and part of the other control network in adults [see Fig. 3; Fair et al., 2007a, 2009]. Thus, while the nature of the task-related BOLD signal did not vary by age or performance, the patterns of functionally stronger inputs and outputs in the dACC/msFC region likely differs over age, and this difference may change (a) the information processing that occurs in this region or (b) information processing in other regions which are functionally related to the region. This point illustrates how information from multiple analysis techniques can inform our understanding of development and provide more nuanced interpretations of the data.

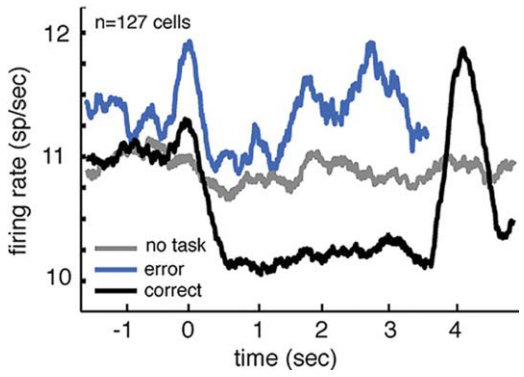

Figure 4.
Neuronal activity in posterior cingulate gyrus (CGp) tracks level of task engagement. This peristimulus time histogram shows average firing rates of a population of 127 cells during error (blue line) and correct (black line) trials of an attentive task, and during the no-task condition (gray line). Neuronal firing rates are suppressed when the macaque monkey is correctly performing the attentive task. Responses on error trials are truncated 500 ms before the error to eliminate any potential perisaccadic or phasic error signals. Adapted from [Hayden et al., 2009].

Negative BOLD Activity

Keeping in mind the neurophysiological constraints of the brain is an important principle for all investigators using human neuroimaging techniques. This principle may be particularly important when comparing groups such as adults and children. While the current spatial resolution of fMRI images necessarily means that we sample from many communities of neurons within a single voxel, it is the parenchymal features that ultimately constrain what can be seen with BOLD contrast-derived signals. Widely recognized by neurophysiologists, but largely unrecognized by neuroimagers, is the fundamental observation that neurons have a baseline rate of firing, and that activity can both increase and decrease from that base rate [Hubel and Wiesel, 1962; Kandel et al., 1991]. The brain is therefore far from silent at rest, and has significant intrinsic activity going on at all times, which may be critical to optimize the way the brain can respond to changing inputs [Chance et al., 2002; Ho and Destexhe, 2000].

When we use BOLD activity measurements to study the brain, we are using an indirect measure of neuronal activity. Though still an area of active investigation, BOLD signal appears to be most related to local field potentials, reflecting primarily the sum of synaptic activity in a volume of neural tissue, and to a lesser extent the spiking (action-potential) outputs [Logothetis, 2003; Logothetis et al., 2001; Mukamel et al., 2005; Raichle and Mintun, 2006]. Postsynaptic potentials (PSPs) comprise the major

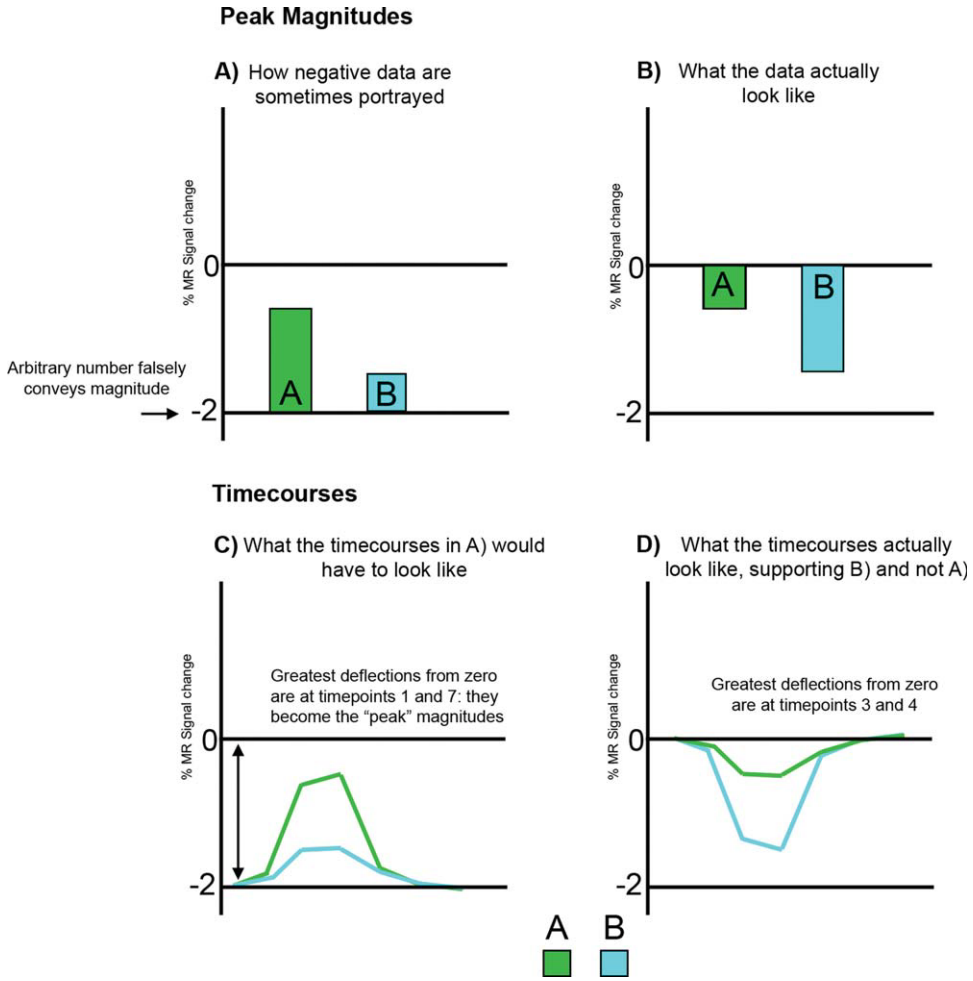

Figure 5.

Negative BOLD activity should be accurately displayed. (**A**) When peak magnitudes are graphed relative to an arbitrary baseline in order to make them appear positive, it erroneously conveys different size magnitudes than are actually present. (**B**) When the peak magnitudes are plotted with bars referenced to a zero baseline, then negative BOLD activity is revealed. (**C**) If the display strategy used in (A) is considered for all timepoints of a timecourse, then the contrived representation of magnitudes is magnified; the greatest difference from zero becomes the initial or final timepoints, which is contrary to the known/measured hemodynamic response function. (**D**) When the data are accurately displayed relative to the baseline, then hypotheses can be appropriately tested.

potentials detected in LFP recordings [Logothetis, 2007]. It logically follows since BOLD activity is related to PSPs, a decrease in PSPs could lead to a decrease in the BOLD signal, and an increase in PSPs could lead to an increase in BOLD signal. The relationship of BOLD to LFPs thus means that BOLD has a greater dynamic range for negative signals than if it was only related to measures of spiking neurons.

And, in fact, there are locations in the brain where cells have a relatively high resting firing rate and then decrease firing as an integral part of a functional network. A classic example is the substantia nigra pars reticulata (SNpr). Inhibitory neurons project from the SNpr to the superior colliculus. Inhibitory activity results in high metabolic needs that could result in an increase in the BOLD response akin to excitatory activity. When the firing rates of these SNpr neurons decrease, saccadic eye movements mediated by the superior colliculus are facilitated [Hikosaka and Wurtz, 1983]. In other words, the decrease in firing rate of neurons in the SNpr inhibits the inhibitory braking effect its efferent projection has on the superior colliculus. In principle, BOLD imaging of the SNpr during a saccade task might reveal decreased signal from baseline, while imaging of the superior colliculus might reveal increased signal. This decrease in firing rate during a task is not just found in basal ganglia and midbrain structures, but also in the cerebral cortex.

Cortical decreases in activity have been recently described in the midline parietal cortex [Hayden et al.,

2009]. Hayden et al. recorded in the presumptive posterior cingulate gyrus in macaques and demonstrated neurons with reliable decreases in firing rate during attention demanding tasks, and relatively increased firing rates during both rest and lapses of attention [see Fig. 4; Hayden et al., 2009]. It is neurobiologically both plausible and likely that this negative dynamic range is played out in PSPs as well. Indeed, as Hayden et al. state:

"Our findings support, and build upon, studies showing that hemodynamic activity in CGp [posterior Cingulate Gyrus] is increased during lapses in attention and failures to perceive and encode environmental stimuli. Moreover, these results show that default effects correspond to spiking activity of CGp neurons, and not just synaptic responses reflecting inputs to CGp. By showing that such effects correspond to the activity of single neurons, and by showing the rapid changes in firing rates associated with task performance, our results significantly advance functional understanding of CGp and, by extension, the default network. More fundamentally, these data confirm the idea that metabolic and hemodynamic changes associated with default processing reflect underlying neurophysiological events and confirm that the default network is homologous in humans and monkeys" [Hayden et al., 2009].

The observation that some neurons, such as those that provide inhibitory afferent tone to their targets, facilitate target activity when they decrease firing rate seems to have been lost in recent times in human neuroimaging. There is a general lack of recognition of negative BOLD activity as something different from positive BOLD activity, and as a result, neurobiologically implausible graphs have recently been published. These graphs often depict apparently negative BOLD change during a condition, but this is never acknowledged and indeed is often disguised with the use of a new, artificial baseline at a negative number instead of zero. This both partially obscures the negative activity (looking at first glance like positive change in BOLD), and mistakenly assumes the traditional zero baseline is arbitrary (see Fig. 5). There is also a fundamental problem where a lack of change between two conditions is interpreted as reflecting the maximum level of processing, but where a negative deflection is interpreted simply as less overall activity in the task of interest. For example, in the case of the SNpr, target-directed saccadic activity would be viewed as overall lower level processing than inter-saccade visual fixation to explain how BOLD would decrease from fixation to saccade in the SNpr. This improbable conceptualization necessitates a floating baseline, forcing negative changes to appear positive. The implausibility of the floating baseline is more obvious when one plots the effect as a timecourse. For the timecourse, the floating baseline must be (1) recalibrated at each moment during the timecourse, and (2) belies the actual estimated shape of the timecourse (see Fig. 5).

Gain and loss of money can be similar to the increase and decrease of BOLD signal. Without having a zero with which to refer to the direction of change, one can be fooled into a misinterpretation of "loss" as a "gain" like two stockbrokers, both having lost money in the stock market, declaring the one of them who has lost less money the "big winner." By acknowledging a baseline level of activity, investigators then are well positioned to use neurophysiological constraints to help contextualize data interpretation when timecourses of BOLD activity are actually less during a task than during rest.

CONCLUSIONS

In conclusion, there are many challenges that developmental researchers face when conducting functional neuroimaging studies. These challenges are not unique to developmental scientists; those investigating cognition across the lifespan as well as those comparing typical and atypical populations face them. While the research question should always drive the experimental approach, to be successful, many different aspects of a study, from acquisition, to atlas alignment, to statistical group comparisons, must be dealt with effectively. Neurophysiologically-constrained interpretation of BOLD signal effects is paramount. Functional neuroimaging provides a remarkable tool to allow us to study cognition across the lifespan and special populations in a safe way. However, while the methods for generating brain images are accessible, that very accessibility can belie the complexity of the work to be done—from initial design to data interpretation.

ACKNOWLEDGMENTS

The authors thank Kelly Anne Barnes and Damien A. Fair for helpful comments on earlier versions of this manuscript.

REFERENCES

Armstrong E, Schleicher A, Omran H, Curtis M, Zilles K (1995): The ontogeny of human gyrification. Cereb Cortex 5:56–63.

Ashburner J, Friston KJ (2000): Voxel-based morphometry—The methods. Neuroimage 11(6, Part 1):805–821.

Biswal B, Yetkin FZ, Haughton VM, Hyde JS (1995): Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. Magn Reson Med 34:537–541.

Brown TT, Lugar HM, Coalson RS, Miezin FM, Petersen SE, Schlaggar BL (2005): Developmental changes in human cerebral functional organization for word generation. Cereb Cortex 15:275–290.

Bunge SA, Dudukovic NM, Thomason ME, Vaidya CJ, Gabrieli JD (2002): Immature frontal lobe contributions to cognitive control in children: Evidence from fMRI. Neuron 33:301–311.

Burgund ED, Kang H-SC, Kelly JE, Buckner RL, Snyder AZ, Petersen SE, Schlaggar BL (2002): The feasibility of a common stereotactic space for children and adults in fMRI studies of development. NeuroImage 17:184–200.

Church JA, Coalson RS, Lugar HM, Petersen SE, Schlaggar BL (2008): A developmental fMRI study of reading and repetition reveals changes in phonological and visual mechanisms over age. Cereb Cortex 18:2054–2065.

Casey BJ (2002): Neuroscience. Windows into the human brain. Science 296:1408–1409.

Caviness VS, Kennedy DN, Richelme C, Rademacher J, Filipek PA (1996): The human brain age 7–11 years: A volumetric analysis based on magnetic resonance images. Cereb Cortex 6:726–736.

Chance FS, Abbott LF, Reyes AD (2002): Gain modulation from background synaptic input. Neuron 35:773–782.

Chang C, Cunningham JP, Glover GH (2009): Influence of heart rate on the BOLD signal: The cardiac response function. Neuroimage 44:857–869.

Church JA, Coalson RS, Lugar HM, Petersen SE, Schlaggar BL (2008): A developmental fMRI study of reading and repetition reveals changes in phonological and visual mechanisms over age. Cereb Cortex 18(9):2054–2065.

Church JA, Fair DA, Dosenbach NU, Cohen AL, Miezin FM, Petersen SE, Schlaggar BL. (2009): Control networks in paediatric Tourette syndrome show immature and anomalous patterns of functional connectivity. Brain 132(Part 1):225–238.

Cohen MS, DuBois RM (1999): Stability, repeatability, and the expression of signal magnitude in functional magnetic resonance imaging. J Magn Reson Imaging 10:33–40.

Corbetta M, Kincade JM, Ollinger JM, McAvoy MP, Shulman GL (2000): Voluntary orienting is dissociated from target detection in human posterior parietal cortex. Nat Neurosci 3:292–297.

Cordes D, Haughton VM, Arfanakis K, Wendt GJ, Turski PA, Moritz CH, Quigley MA, Meyerand ME (2000): Mapping functionally related regions of brain with functional connectivity MR imaging. AJNR Am J Neuroradiol 21:1636–1644.

Damoiseaux JS, Rombouts SA, Barkhof F, Scheltens P, Stam CJ, Smith SM, Beckmann CF (2006): Consistent resting-state networks across healthy subjects. Proc Natl Acad Sci USA 103: 13848–13853.

Dosenbach NU, Visscher KM, Palmer ED, Miezin FM, Wenger KK, Kang HC, Burgund ED, Grimes AL, Schlaggar BL, Petersen SE (2006): A core system for the implementation of task sets. Neuron 50:799–812.

Dosenbach NU, Fair DA, Miezin FM, Cohen AL, Wenger KK, Dosenbach RAT, Fox MD, Snyder AZ, Vincent JL, Raichle ME, Schlaggar BL, Petersen SE (2007): Distinct brain networks for adaptive and stable task control in humans. Proc Natl Acad Sci USA 104:11073–11078.

Dosenbach NU, Fair DA, Cohen AL, Schlaggar BL, Petersen SE (2008): A dual-networks architecture of top-down control. TICS 12:99–105.

Fair DA, Brown TT, Petersen SE, Schlaggar BL (2006): A comparison of ANOVA and correlation methods for investigating cognitive development with fMRI. Dev Neuropsychol 30:531–546.

Fair DA, Dosenbach NUF, Church JA, Cohen AL, Brahmbhatt S, Miezin FM, Barch DM, Raichle ME, Petersen SE, Schlaggar BL (2007a): Development of distinct control networks through segregation and integration. Proc Natl Acad Sci USA 104:13507–13512.

Fair DA, Schlaggar BL, Cohen AL, Miezin FM, Dosenbach NU, Wenger KK, Fox MD, Snyder AZ, Raichle ME, Petersen SE (2007b): A method for using blocked and event-related fMRI data to study "resting state" functional connectivity. Neuroimage 35:396–405.

Fair DA, Cohen AL, Dosenbach NU, Church JA, Miezin FM, Barch DM, Raichle ME, Petersen SE, Schlaggar BL (2008): The maturing architecture of the brain's default network. Proc Natl Acad Sci USA 105:4028–4032.

Fair DA, Cohen AL, Power JD, Dosenbach NUF, Church JA, Miezin FM, Schlaggar BL, Petersen SE (2009): Functional brain networks develop from a "local to distributed" organization. PLoS Comput Biol 5:e1000381.

Fox MD, Raichle ME (2007): Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. Nat Rev Neurosci 8:700–711.

Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, Raichle ME (2005): The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proc Natl Acad Sci USA 102:9673–9678.

Friston KJ, Price CJ, Fletcher P, Moore C, Frackowiak RSJ, Dolan RJ (1996): The trouble with cognitive subtraction. NeuroImage 4:97–104.

Garavan H, Ross TJ, Murphy K, Roche RA, Stein EA (2002): Dissociable executive functions in the dynamic control of behavior: Inhibition, error detection, and correction. Neuroimage 17:1820–1829.

Giedd JN, Blumenthal J, Jeffries NO, Castellanos FX, Liu H, Zijdenbos A, Paus T, Evans AC, Rapoport JL (1999): Brain development during childhood and adolescence: A longitudinal MRI study. Nat Neurosci 2:861–863.

Grossman M, Cooke A, DeVita C, Alsop D, Detre J, Chen W, Gee J (2002): Age-related changes in working memory during sentence comprehension: An fMRI study. Neuroimage 15:302–317.

Hayden BY, Smith DV, Platt ML (2009): Electrophysiological correlates of default-mode processing in macaque posterior cingulate cortex. Proc Natl Acad Sci USA 106:5948–5953.

Hikosaka O, Wurtz RH (1983): Visual and oculomotor functions of monkey substantia nigra pars reticulata. I. Relation of visual and auditory responses to saccades. J Neurophysiol 49:1230–1253.

Ho N, Destexhe A (2000): Synaptic background activity enhances the responsiveness of neocortical pyramidal neurons. J Neurophysiol 84:1488–1496.

Honey C, Kotter R, Breakspear M, Sporns O (2007): Network structure of cerebral cortex shapes functional connectivity on multiple time scales. Proc Natl Acad Sci USA 104:10240–10245.

Hubel DH, Wiesel TN (1962): Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J Physiol (London) 160:106–154.

Ihnen SKZ, Church JA, Petersen SE, Schlaggar BL (2009): Lack of generalizability of sex differences in the fMRI BOLD activity associated with language processing in adults. Neuroimage 45:1020–1032.

Johnson MH, Halit H, Grice SJ, Karmiloff-Smith A (2002): Neuroimaging of typical and atypical development: A perspective from multiple levels of analysis. Dev Psychopathol 14:521–536.

Jones TB, Bandettini PA, Kenworthy L, Case LK, Milleville SC, Martin A, Birn RM (2010): Sources of group differences in functional connectivity: An investigation applied to autism spectrum disorder. Neuroimage 49:401–414.

Just MA, Cherkassky VL, Keller TA, Kana RK, Minshew NJ (2007): Functional and anatomical cortical underconnectivity in autism: Evidence from an FMRI study of an executive function task and corpus callosum morphometry. Cereb Cortex 17:951–961.

Kail R (1991): Developmental change in speed of processing during childhood and adolescence. Psychol Bull 109:490–501.

Kandel ER, Schwartz JH, Jessel TM (1991): Principles of Neural Science.New York:Elsevier Science.1135 p.

Kang HC, Burgund ED, Lugar HM, Petersen SE, Schlaggar BL (2003): Comparison of functional activation foci in children and adults using a common stereotactic space. Neuroimage 19:16–28.

Karrer B, Levina E, Newman ME (2008): Robustness of community structure in networks. Phys Rev E Stat Nonlin Soft Matter Phys 77(4, Part 2):046119.

Kelly AM, Margulies DS, Castellanos FX. (2007): Recent advances in structural and functional brain imaging studies of attention-deficit/hyperactivity disorder. Curr Psychiatry Rep 9:401–407.

Konrad K, Neufang S, Thiel CM, Specht K, Hanisch C, Fan J, Herpertz-Dahlmann B, Fink GR (2005): Development of attentional networks: An fMRI study with children and adults. Neuroimage 28:429–439.

Kotsoni E, Byrd D, Casey BJ (2006): Special considerations for functional magnetic resonance imaging of pediatric populations. J Magn Resonance Imaging 23:877–886.

Lenroot RK, Giedd JN (2006): Brain development in children and adolescents: Insights from anatomical magnetic resonance imaging. Neurosci Biobehav Rev 30:718–729.

Logothetis NK (2003): The underpinnings of the BOLD functional magnetic resonance imaging signal. J Neurosci 23:3963–3971.

Logothetis NK (2007): The ins and outs of fMRI signals. Nat Neurosci 10:1230–1232.

Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A (2001): Neurophysiological investigation of the basis of the fMRI signal. Nature 412:150–157.

Lowe MJ, Mock BJ, Sorenson JA (1998): Functional connectivity in single and multislice echoplanar imaging using resting-state fluctuations. Neuroimage 7:119–132.

Mathur AM, Neil JJ, McKinstry RC, Inder TE (2008): Transport, monitoring, and successful brain MR imaging in unsedated neonates. Pediatr Radiol 38:260–264.

Mukamel R, Gelbard H, Arieli A, Hasson U, Fried I, Malach R (2005): Coupling between neuronal firing, field potentials, and FMRI in human auditory cortex. Science 309:951–954.

Mumford JA, Nichols TE (2008): Power calculation for group fMRI studies accounting for arbitrary design and temporal autocorrelation. Neuroimage 39:261–268.

O'Shaughnessy ES, Berl MM, Moore EN, Gaillard WD (2008): Pediatric functional magnetic resonance imaging (fMRI): Issues and applications. J Child Neurol 23:791–801.

Palmer ED, Brown TT, Petersen SE, Schlaggar BL (2004): Investigation of the functional neuroanatomy of single word reading and its development. Sci Stud Read 8:203–223.

Posner MI (1978): Chronometric Explorations of Mind. Englewood Heights, NJ: Lawrence Erlbaum Associates.

Price CJ, Friston KJ (1999): Scanning patients with tasks they can perform. Hum Brain Mapp 8:102–108.

Raichle ME, Mintun MA (2006): Brain work and brain imaging. Annu Rev Neurosci 29:449–476.

Raschle NM, Lee M, Buechler R, Christodoulou JA, Chang M, Vakil M, Stering PL, Gaab N (2009): Making MR imaging child's play—Pediatric neuroimaging protocol, guidelines and procedure. JoVE 29, http://www.jove.com/index/details.stp?id1309, doi: 10.3791/1309.

Reuter-Lorenz PA, Lustig C (2005): Brain aging: reorganizing discoveries about the aging mind. Curr Opin Neurobiol 15:245–251.

Rubia K, Smith AB, Brammer MJ, Toone B, Taylor E (2005): Abnormal brain activation during inhibition and error detection in medication-naive adolescents with ADHD. Am J Psychiatry 162:1067–1075.

Schlaggar BL, McCandliss BD (2007): Development of neural systems for reading. Annu Rev Neurosci 30:475–503.

Schlaggar BL, Brown TT, Lugar HM, Visscher KM, Miezin FM, Petersen SE (2002): Functional neuroanatomical differences between adults and school-age children in the processing of single words. Science 296:1476–1479.

Shulman GL, Ollinger JM, Akbudak E, Conturo TE, Snyder AZ, Petersen SE, Corbetta M (1999): Areas involved in encoding and applying directional expectations to moving objects. J Neurosci 19:9480–9496.

Thirion B, Pinel P, Meriaux S, Roche A, Dehaene S, Poline J-B (2007): Analysis of a large fMRI cohort: Statistical and methodological issues for group analyses. Neuroimage 35:105–120.

Thomason ME, Burrows BE, Gabrieli JD, Glover GH (2005): Breath holding reveals differences in fMRI BOLD signal in children and adults. Neuroimage 25:824–837.

Velanova K, Wheeler ME, Luna B (2008): Maturational changes in anterior cingulate and frontoparietal recruitment support the development of error processing and inhibitory control. Cereb Cortex 18:2505–2522.

Velanova K, Wheeler ME, Luna B (2009): The maturation of task set-related activation supports late developmental improvements in inhibitory control. J Neurosci 29:12558–12567.

Yerys BE, Jankowski KF, Shook D, Rosenberger LR, Barnes KA, Berl MM, Ritzl EK, VanMeter J, Vaidya CJ, Gaillard WD (2009): The fMRI success rate of children and adolescents: Typical development, epilepsy, attention deficit/hyperactivity disorder, and autism spectrum disorders. Hum Brain Mapp 30:3426–3435.

Yuan W, Altaye M, Ret JR, Schmithorst VJ, Byars AW, Plante E, Holland SK (2009): Quantification of head motion in children during various fMRI language tasks. Hum Brain Mapp 30:1481–1489.